

# Abstracts

## Oral 15

## Data analysis

### ○15.1 THE APPLICATION OF STATISTICAL TESTS TO CLINICAL DIAGNOSTICS

M. Boggess<sup>1</sup>, M. Guest<sup>2</sup>, J. Attia<sup>2</sup>, C. D'Este<sup>2</sup>, A. Brown<sup>3</sup>. <sup>1</sup>Stata Corporation, College Station, TX, USA; <sup>2</sup>Faculty of Health, University of Newcastle, Callaghan, NSW, Australia; <sup>3</sup>Population Health, Macquarie Area Health Service, Dubbo, NSW, Australia

**Introduction:** Medical diagnostic tests are developed by clinicians to obtain reliable diagnoses. Thus not all diagnostic tests give results that are amenable to statistical examination. This poses a particular problem for epidemiologists: in order to be able find significant evidence of an effect of exposure we need to use a sample size appropriate for the statistical test being used. This may require examining the test results of many subjects.

**Methods:** A good example of an intractable diagnostic test is the L'Anthony Desaturated Panel-D-15 for colour vision deficiency. An ophthalmologist makes a diagnosis, into one of several categories, by examining a diagram that is a representation of the results of the test. In the 1980s two attempts (Bowman and Vingrys *et al*) were made to compute a numerical index from the L'Anthony test results. While this made it possible to test hypotheses concerning exposure, the indices are difficult to interpret and it has been unclear as to the distributional properties of these indices. We examine the calculation of these indexes in the light of statistical data reduction techniques.

**Results:** We show that Vingrys' "moment of inertia" method is essentially principle component analysis, a data reduction method which uses matrix methods, commonly used in social sciences. Using data from the Australian Aircraft Maintenance Technician study, we apply discriminant analysis to derive a decision rule for diagnosis, which we compare to current methods using ROC curves.

### ○15.2 INVESTIGATING COLOUR VISION IMPAIRMENT IN AIRCRAFT MAINTENANCE TECHNICIANS

M. Guest<sup>1</sup>, J. Attia<sup>1</sup>, C. D'Este<sup>1</sup>, A. Brown<sup>2</sup>, M. Boggess<sup>3</sup>. <sup>1</sup>Faculty of Health, University of Newcastle, Callaghan, NSW, Australia; <sup>2</sup>Population Health, Macquarie Area Health Service, Dubbo, NSW, Australia; <sup>3</sup>Stata Corporation, College Station, TX, USA

**Introduction:** In order to determine if there was evidence to support the anecdotal reports of adverse health problems, including deficiencies in colour vision, in aircraft maintenance technicians who undertook deseal/reseal activities on the Australian Air Force's F-111 aircraft at Amberley Air Base, a general health and medical study was undertaken. The deseal/reseal process required the personnel to enter the fuel tanks where there were excessive exposures to formulations containing solvents such as MEK, toluene (aromatic naptha), thiophenol, and propylene glycol (monomethyl ether acetate, as well as primers and sealants containing chromates, unreacted isocyanates, and curing agents.

**Methods:** Colour vision was assessed in each eye of 616 exposed personnel, 516 unexposed personnel from similar technical trades and 406 non-technical comparisons using the L'Anthony Desaturated Panel-D-15. The results were assessed by plotting cap arrangement on a circular chart to provide a clinical diagnosis, and calculation of Bowman's Colour Confusion Index (CCI). As the distribution of CCI required the use of a non-linear model, Vingrys's C-index, S-index, and angle of confusion were also calculated.

**Results:** Over half of the study participants were diagnosed with a colour vision deficit. Multivariate regression analysis of the dichotomised CCI (normal v non-normal) controlling for confounders including age, alcohol consumption, smoking, diabetes, and visual acuity, indicated a borderline increase in abnormal colour vision in the exposed population. The increase became significant in the reduced model (OR 1.37; 95% CI 1.02 to 1.85).

### ○15.3 IMPACT OF UNCERTAINTIES ASSOCIATED WITH RADON EXPOSURE IN DOSE-RESPONSE ANALYSES AMONG URANIUM MINERS

A. Rogel<sup>1</sup>, V. Erzen<sup>1</sup>, M. Tirmarche<sup>1</sup>, D. Laurier<sup>1</sup>. <sup>1</sup>Institute for Radiological Protection and Nuclear Safety (IRSN), France

**Introduction:** The work is motivated by the need to estimate the impact of exposure uncertainties in occupational dose-response analyses.

**Methods:** The French uranium miners' cohort includes 5098 men working in mines since 1946. The aim of the study is to estimate the risk of lung cancer associated to cumulated radon exposure. Quantitative annual exposures were known for all members of the cohort. However, in the first 10 years of follow up, exposures were retrospectively estimated based on job title and on environmental measures, whereas in the recent years exposures were directly measured using dosimeters. Scenarios were simulated in which annual values of radon exposure were replaced by random numbers generated according to a gamma or uniform distribution. Impact on dose-response and dose-time-response analyses was evaluated.

**Results:** Lung cancer risk associated to radon exposure was always significant whatever the hypothesis. Estimated relative risk varied by a factor 3 according to exposure scenario. Nonetheless a slight underestimation of the dose-response due to exposure uncertainties during the old years was suspected. Relative risk decreased with time since exposure in most scenarios, but this modifying effect is much more sensitive to hypotheses on exposure uncertainties.

**Conclusion:** An evaluation of the impact of uncertainties is a necessary step of the dose-response analysis. The proposed methodology can be applied to other occupational cohorts where exposure is quantitatively reconstructed on annual basis or where reconstruction quality varies with time.

### ○15.4 ON THE LABILITY OF ODDS RATIOS

D. C. Glass<sup>1</sup>, C. N. Gray<sup>2</sup>, D. J. Jolley<sup>2</sup>, C. Gibbons<sup>2</sup>, M. R. Sim<sup>1</sup>. <sup>1</sup>Monash University, Melbourne, Australia. <sup>2</sup>Deakin University Australia

**Introduction:** The Health Watch case control study examined the relation between benzene exposure and lympho-haematopoietic disorders, specifically leukaemia, in the petroleum industry. This paper examines the effect on the calculated odds ratios (ORs) of different ways of analysing the data and of measuring exposure.

**Methods:** The cumulative exposure to benzene was expressed in ppm-years and the ORs were calculated by logistic regression using benzene exposure as a continuous variable and also as a categorical variable.

**Results:** Unmatched analysis resulted in a lower OR 35.8 (95% CI 6.8 to 189) than did matched analysis, OR 98.2 (8.8-1090) for the highest exposure category (>16 ppm-years) compared to the lowest exposure category (<1 ppm-years). Combining the two lowest exposure categories to form a new reference category (<2 ppm-years) reduced the matched OR to 51.9 (5.6 to 480). When the cutpoint for the highest exposure category was changed to match those used in other similar petroleum industry studies (compared to <1 ppm-years category) the matched ORs were again reduced: a cutpoint of 4.8 ppm-years resulted in an OR 2.5 (1.1 to 5.7); while a cut-point at 8 ppm-years resulted in an OR of 11.3 (2.8 to 45.1). We investigated which subjects might have had accidental exposures and added these to the cumulative exposure estimates. The increases were mainly attributed to the more highly exposed workers. The OR was reduced to 7.8 (2.3 to 25.9) for the highest exposure category. A similar reduction was observed when exposure was treated as a continuous variable and accidental exposures included. This reduction in ORs is probably because the leukemia risk is associated with higher exposures when the accidental exposures are added and hence the risk per ppm-year is reduced.

**Conclusions:** Various analyses were performed which showed that there was a strong relation between benzene exposure and risk of leukaemia. The observed odds ratios for the highest exposure category in the same data set ranged from 98 to 2.5 depending on the type of analysis, the exposures included, the choice of reference group, and cutpoint. However, the lower confidence intervals were greater than 1 in all cases, hence these ORs should be regarded as a strong indication of causality according to Bradford Hill's first criterion of strength of association.

**O15.5 TWO-STAGE METHODS FOR POPULATION BASED STUDIES: INTRODUCING MULTIPLE EXPOSURE CATEGORIES AND CONFOUNDING FACTORS**

A. Oudin, J. Björk, U. Strömberg. *Department of Occupational and Environmental Medicine, Lund University, Sweden*

**Introduction:** Two-stage methods can be attractive. In the first stage, register data are retrieved for all study subjects. The subjects can be grouped to occupational group or residential area; group level information on exposure is obtained, besides general individual characteristics, such as age and sex. In the second stage, individual data are obtained for the subjects who participated in an interview/questionnaire investigation. The expectation-maximisation (EM) method has been suggested for effect estimation based on first and second stage data, considering a dichotomous exposure (Strömberg & Björk. *Epidemiology* 2004;15:494–503). Here, we consider more general situations, focusing on a planned study on air pollution and stroke in southern Sweden.

**Methods:** We consider a two-stage case control study, collecting all stroke cases incident during a year in 12 municipalities in Skåne, southern Sweden. Controls are selected from the same population.

Questionnaire data are collected for the second stage subjects. Three levels of exposure are introduced; low, medium, and high. Furthermore, smoking is a dichotomous confounder. Three methods are used for effect estimation: (1) second stage data only; (2) EM method where second stage data are used to calculate the group level exposure probabilities; and (3) EM method where group level exposure probabilities are obtained from register data. No bias is introduced. The confidence intervals for the EM method are estimated with a bootstrapping technique.

**Results:** All estimates are bias free and the coverages are around 95%. The precision of the estimates increases with the EM method. The lowest standard errors are obtained when using the EM method with register data. If register data are erroneous, the EM method that uses register data may yield biased estimates.

**Conclusion:** The precision can increase by using the EM method. We recommend estimating the effect with all three methods. If the EM estimates differ, register data might be erroneous, if the EM estimates are similar; the gain in precision is largest with the EM method that uses register data. If the estimates obtained with second stage data differ from the estimates obtained with the EM methods, participation bias might be present.